Ames
Research
Center

Accurate Biomass Estimation via Bayesian Adaptive Sampling

Kevin R. Wheeler, NASA ARC

Kevin H. Knuth, University at Albany

Joseph P. Castle, EAP at NASA ARC

Nikolay Lvov, QSS @ NASA ARC

kwheeler@mail.arc.nasa.gov

Outline

- ***Problem*** – Quantifying uncertainty on biomass estimations.
- ***Approach*** – Bayesian adaptive sampling with robotic platforms.
- ***Intermediate Results*** – Analysis of MISR data.
- ***Future Work***

Biomass Estimation Problem

Ames Research Center

How much wood is standing in the forest – related to carbon sequestration and and CO2 sink estimation.

Bottom-Up Approach – USFS measures diameter at breast height (dbh) and tables of allometry to estimate carbon (e.g. BIOPAK). Finite & fixed sampling approach. No PDFs.

Top-Down Approach – Remote sensing estimates with MODIS & MISR for NPP, LAI, BHR, land cover campaigns. Poor spatial resolution, scaling problems, unquantified uncertainty.

Two approaches for same problem but:

1. not integrated
2. no single unifying model
3. missing intermediate "*ground truth*" sources
4. many data sources are available that are not incorporated
5. no comparable error statistics
6. predetermined LTERs do not maximize uncertainty reduction.

Biomass Estimation
Data Sources

Ground *Truth*

* diameter at breast height with allometry tables
* measurements with PARABOLA.
* LTERs

Various airborne sorties

Remote Sensing –

* MODIS IGBP - landcover maps – 1 km
* MODIS - global products (NPP…) – 250m – 1 km
* MISR - global products (LAI, FPAR) – 1km
* QuickBird – photos
* SAR
* LIDAR
* iKONOS – photos
* Landsat 7

 kwheeler@mail.arc.nasa.gov

Biomass Estimation
What we want to know

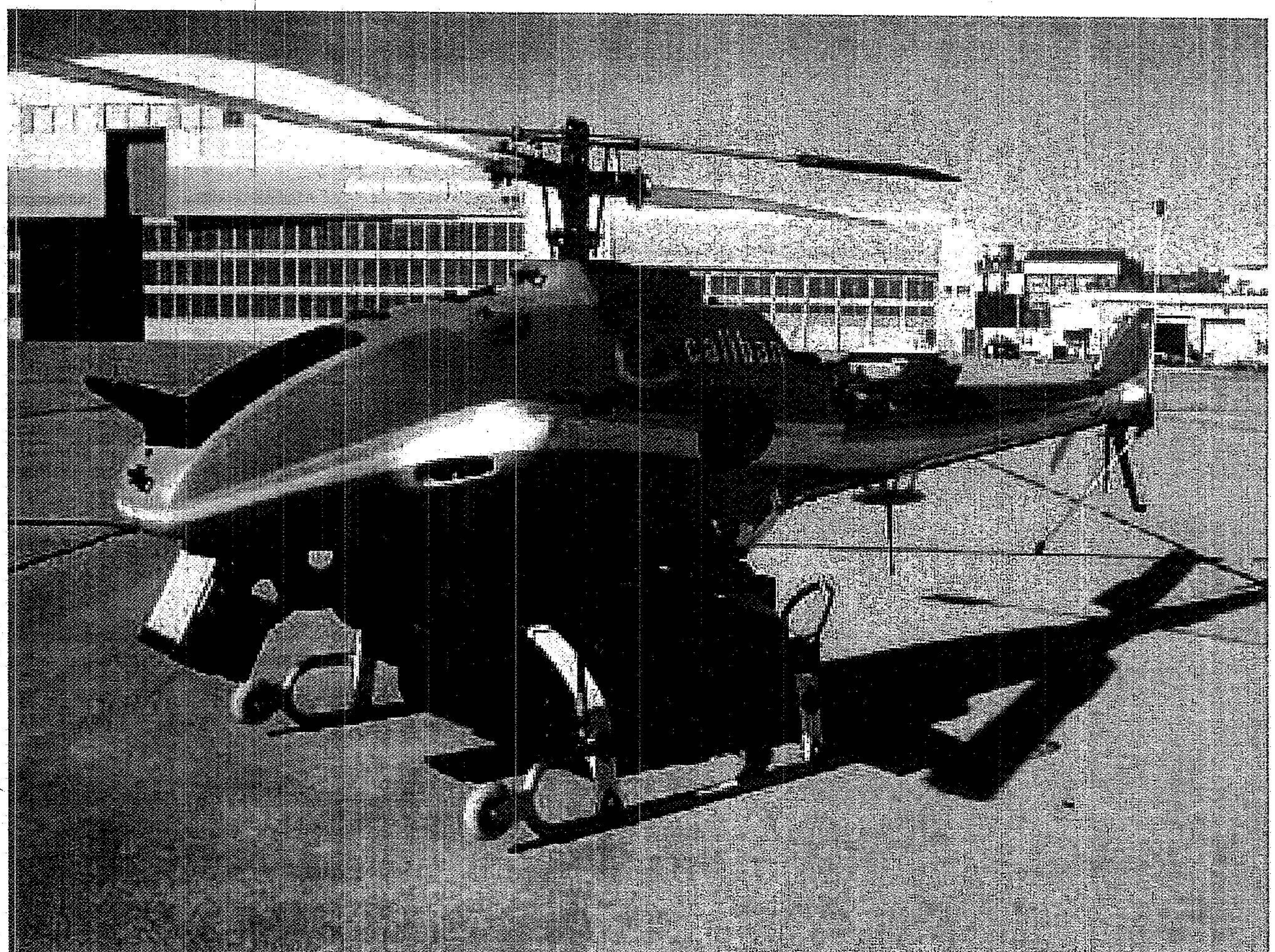
Ideally would like to know of each tree on the planet:

- base diameter
- canopy diameter
- height
- location
- species

4 out of 5 is not bad…

kwheeler@mail.arc.nasa.gov

Automated Sampling Platform



NASA ARC & US Army
(Matthew Whalley)

Autonomous rotocraft

- LIDAR
- Spectrometer (Viz-NIR)
- Stereo Cameras
- Differential GPS
- Accurate IMU
- Wireless Ethernet

Will autonomously fly
to specified way-points.

kwheeler@mail.arc.nasa.gov

Ames Research Center

Automated Sampling Platform Problem

Problems:

1. Ideally would like to cover at least 1×10^6 Km^2
2. Want hemispherical sampling for each ground/biome patch.
3. Want sampling at multiple illumination angles.

Answer:

1. Only sample where you need to.
2. Combine with other data sources.
3. Use a common model.

Bayesian Adaptive Sampling

Goal: Sample with respect to maximally reducing uncertainty in a representative model's posterior.

Definition:

$p(\text{model}) =$ probability of the model parameters taking on particular values

Bayes Posterior:

$$p(\text{model} \mid \text{data}) = \frac{p(\text{data} \mid \text{model})\, p(\text{model})}{\int p(\text{data} \mid \text{model})\, p(\text{model})\, d(\text{model})}$$

kwheeler@mail.arc.nasa.gov

$E\{I(\gamma)\} = -\int d\gamma p(\gamma \mid D)\log[p(\gamma \mid D)]$

Ames Research Center

BAS: Maximizing uncertainty reduction

Want to select in a rigorous manner the next observation which will give us the most information. Shannon showed us that expected information is represented by the negative entropy:

$$E\{I(\gamma)\} = -\int d\gamma\, p(\gamma \mid D)\log[p(\gamma \mid D)]$$

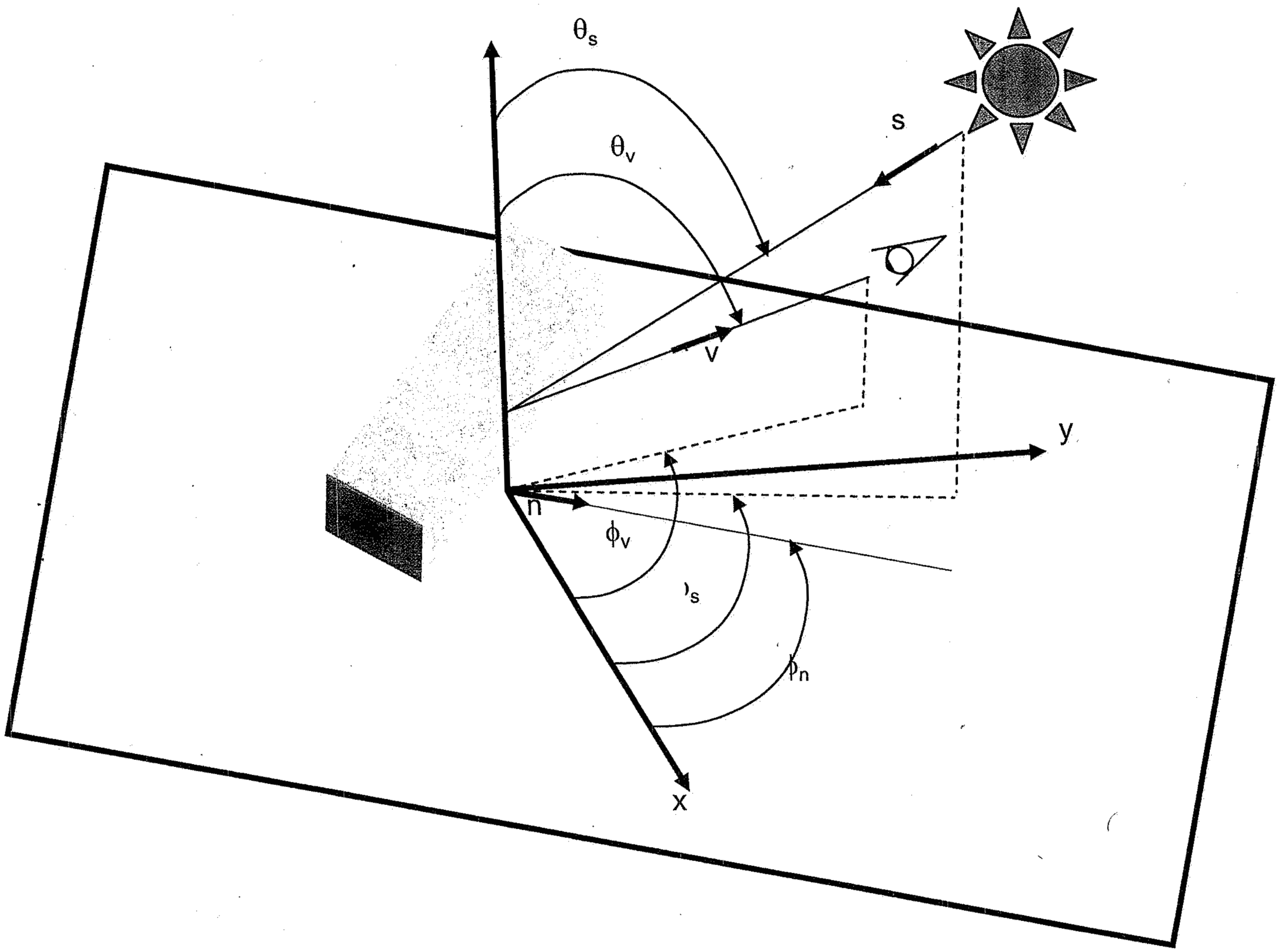
The entropy of the posterior tells us how much information we currently have about our model parameters.

The change in the entropy of the predictive posterior from one observation to the next tells us how much information we have gained/lost with that sample. We use Monte Carlo simulations to sample the predictive distribution.

The next observation that we should take is the one that maximizes this information gain and satisfies the platform constraints.

Ames
Research
Center

BAS: BRDF



kwheeler@mail.arc.nasa.gov

BAS: Forward Model

Ames Research Center

MISR Linearized Rahman model:

$$\rho_{\mathrm{MOD}}(-\mu_s,\mu_v,\phi_s,\phi_v) = r_0 \frac{\mu_s^{k-1}\mu_0^{k-1}}{(\mu_s+\mu_v)^{1-k}} \exp\left(b\cdot p(\Omega)\right)\cdot h(-\mu_s,\mu_v,\phi_s-\phi_v)$$

Free parameters: r_0 – *reflectance*, b – *scattering*, k – *slope term*

where

$$\mu_s = \cos(\theta_s) \qquad \mu_v = \cos(\theta_v)$$

Scattering angle:

$$p(\Omega) = \cos(\Omega) = -\mu_s\mu_v + \sqrt{1-\mu_s^2}\cdot\sqrt{1-\mu_v^2}\cdot\cos(\phi_s-\phi_v)$$

Hot spot:

$$h(-\mu_s,\mu_v,\phi_s-\phi_v) = 1 + \frac{1-r_0}{1+G(-\mu_s,\mu_v,\phi_s-\phi_v)}$$

$$G(-\mu_s,\mu_v,\phi_s-\phi_v) = \left\{\left(\frac{1}{\mu_s^{\,2}}-1\right)+\left(\frac{1}{\mu_v^2}-1\right)+2\left[\sqrt{\left(\frac{1}{\mu_s^2}-1\right)\left(\frac{1}{\mu_v^2}-1\right)}\right]\cos(\phi_s-\phi_v)\right\}^{1/2}$$

kwheeler@mail.arc.nasa.gov

Ames Research Center

BAS: Forward Model

Mixed biome model:

$$\rho(\theta_s,\phi_s;\theta_v,\phi_v,\lambda)=\sum_{i=1}^{6} w_i\, p(\theta_s,\phi_s;\theta_v,\phi_v;r_{0,i,\lambda},k_{i,\lambda},b_{i,\lambda},\lambda)$$

Treat the following as unobserved random variables:

- weights of biomes within a mixed pixel (w_b)
- noise on angles
- for each biome and wavelength:
- r_0 – reflectance
- k – slope
- b - scattering

kwheeler@mail.arc.nasa.gov

Ames
Research
Center

MISR Rahman Model Free Parameters Statistics

- Separable in 12 dimensions (4 bands x 3 free parameters)?
- MODIS Landcover: 6 biomes: grasses & crops, shrubs, broadleaf crops, savannas, broadleaf forests, and needle leaf forests.
- Mixed biomes in 1 Km pixels.
- MISR data conditioned upon MODIS IGBP land cover (85% threshold) using published free parameters and output.

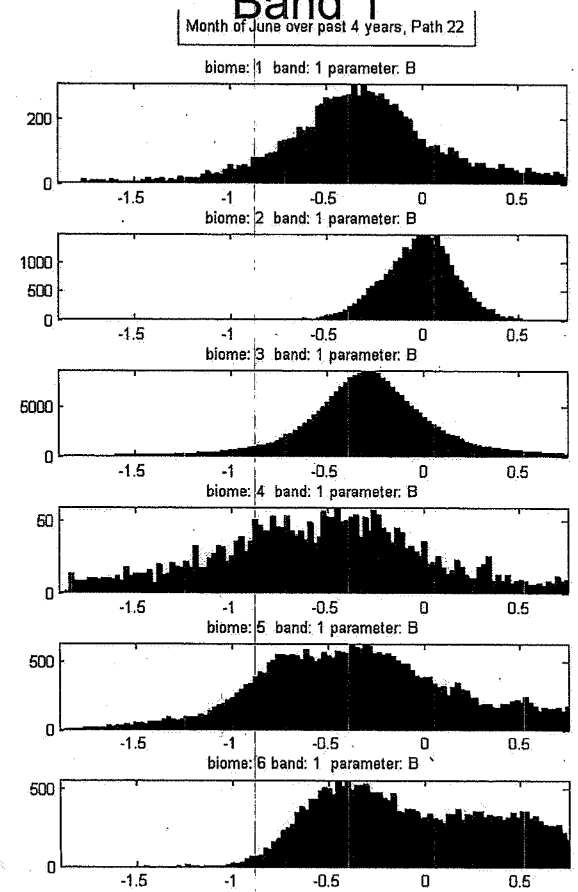
kwheeler@mail.arc.nasa.gov

MISR Rahman Model
Free Parameters Histograms

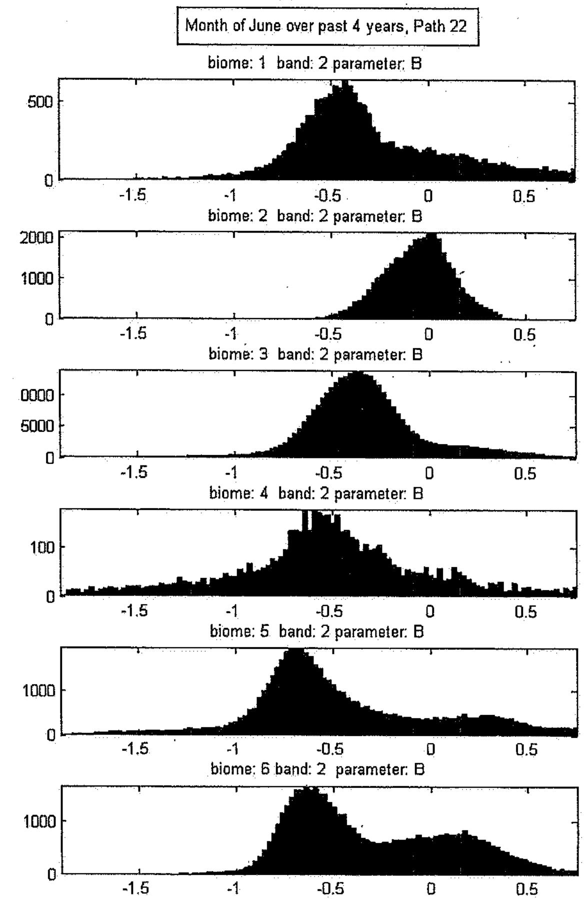
Ames
Research
Center

Parameter B (scattering)
Month of June, 2005 – 2001, Path 22
6 MODIS biomes: grasses & crops, shrubs, broadleaf crops, savannas, broadleaf forests, and needle leaf forests

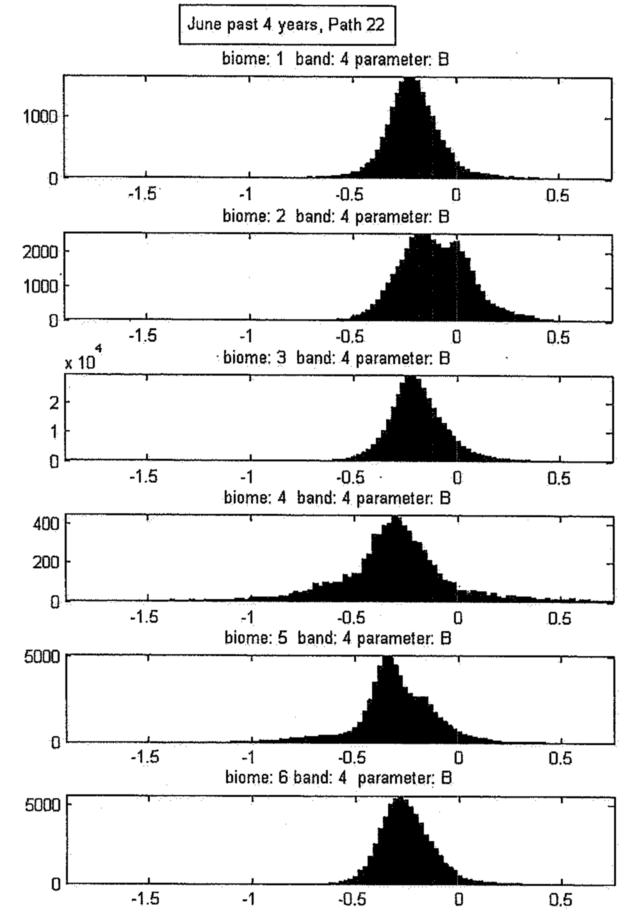
Band 1



Band 2



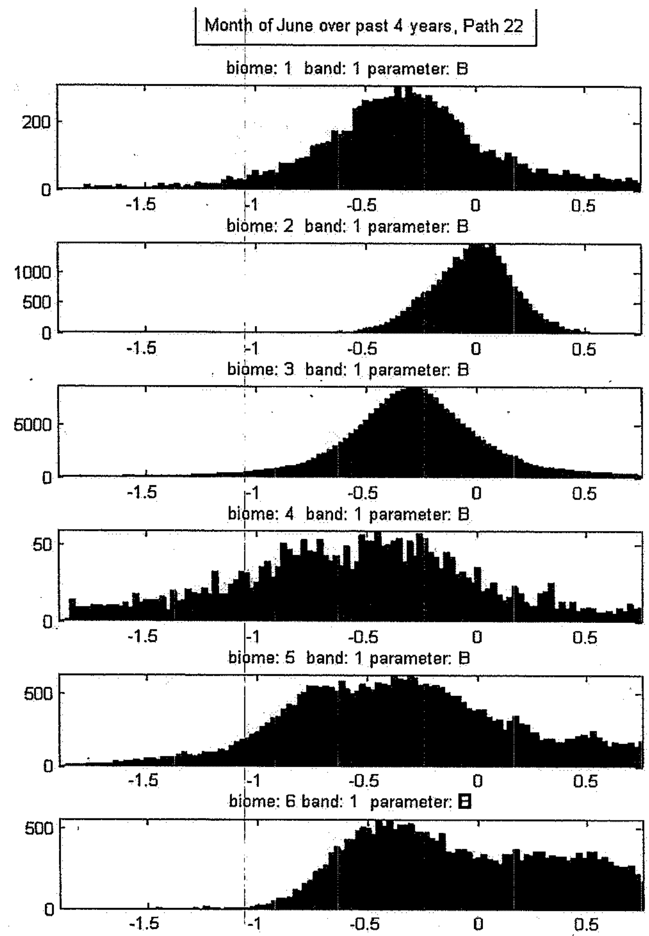
Band 4



MISR Rahman Model
Free Parameters Histograms

Parameter B (scattering), Band 1 (Red)
Month of June and January, 2005 – 2001, Path 22
6 MODIS biomes: grasses & crops, shrubs, broadleaf crops, savannas, broadleaf forests, and needle leaf forests

June



Grasses & crops

Shrubs

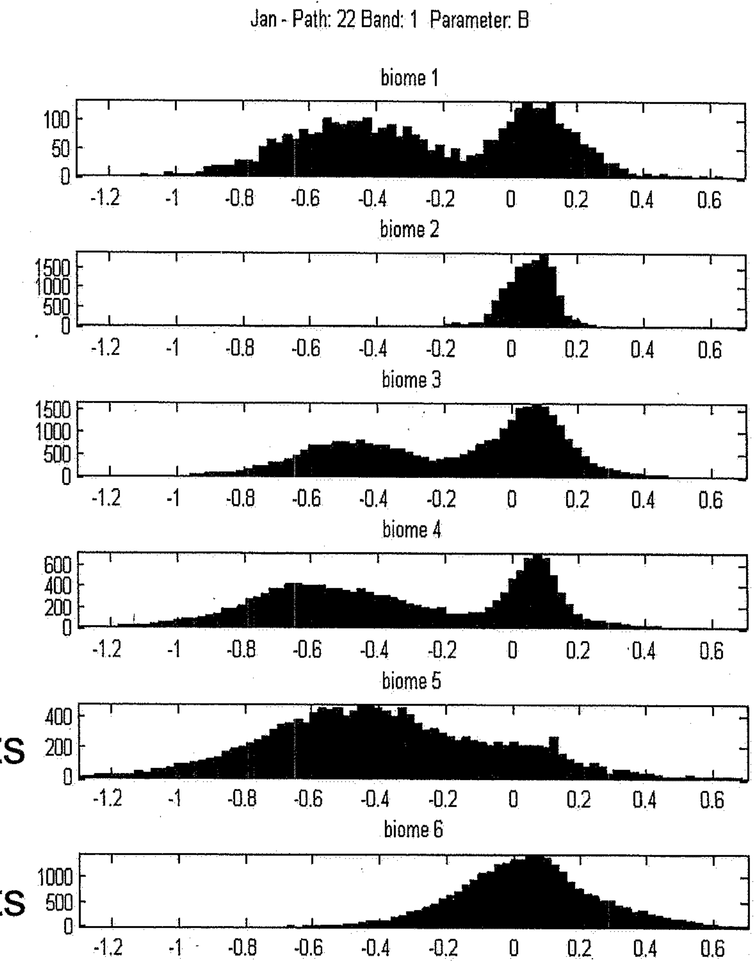
Broadleaf crops

Savannas

Broadleaf Forests

Needle leaf forests

January

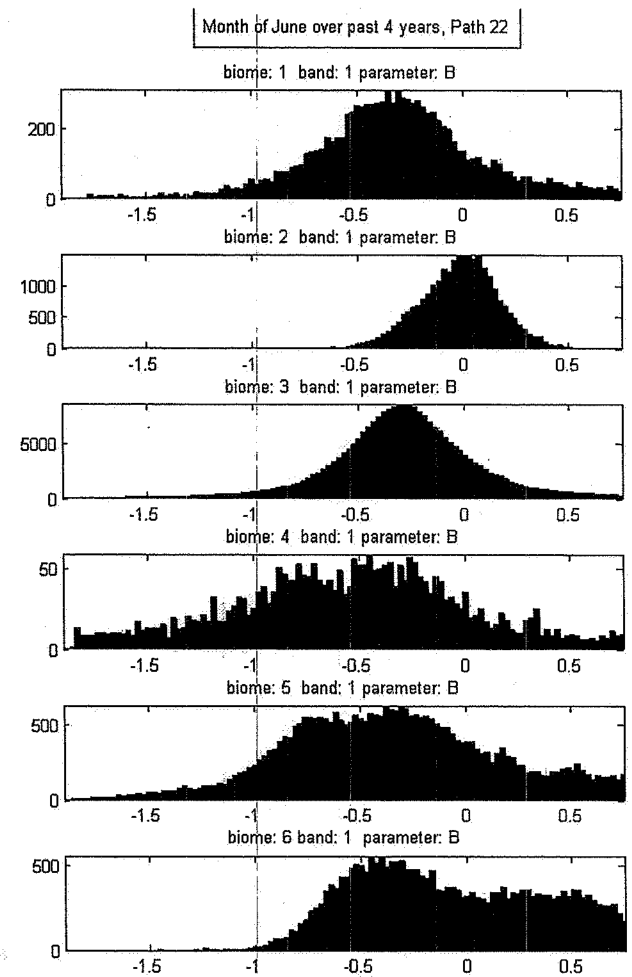


vheeler@mail.arc.r

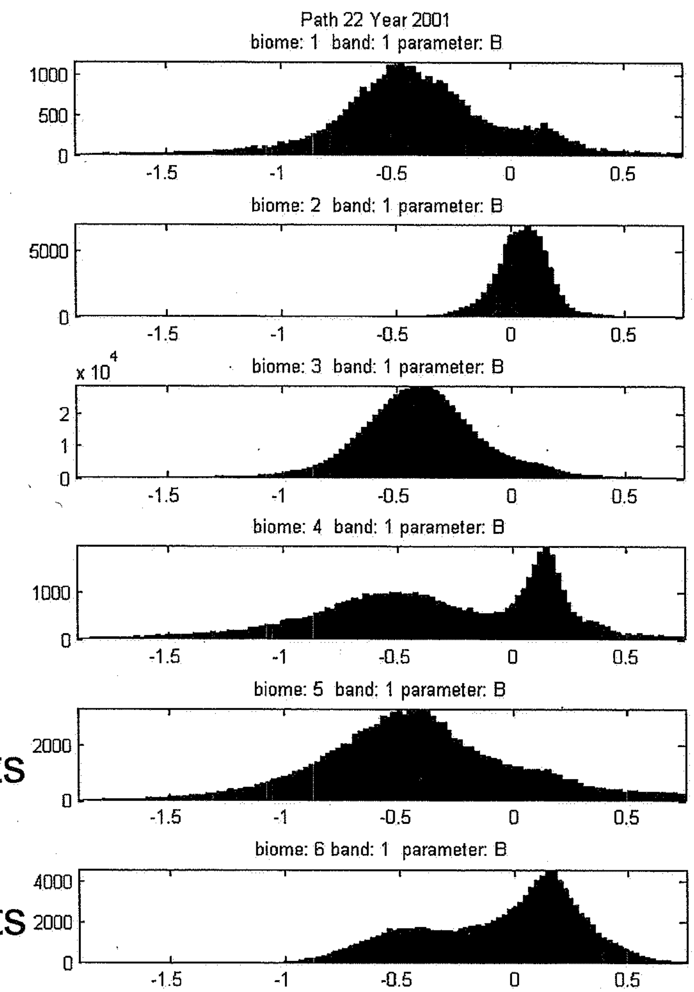
Ames
Research
Center

MISR Rahman Model
Free Parameters Histograms

Path 22, June, Band 1, Param. B



Path 22, All years, Band 1, Param B.



Grasses & crops

Shrubs

Broadleaf crops

Savannas

Broadleaf Forests

Needle leaf forests

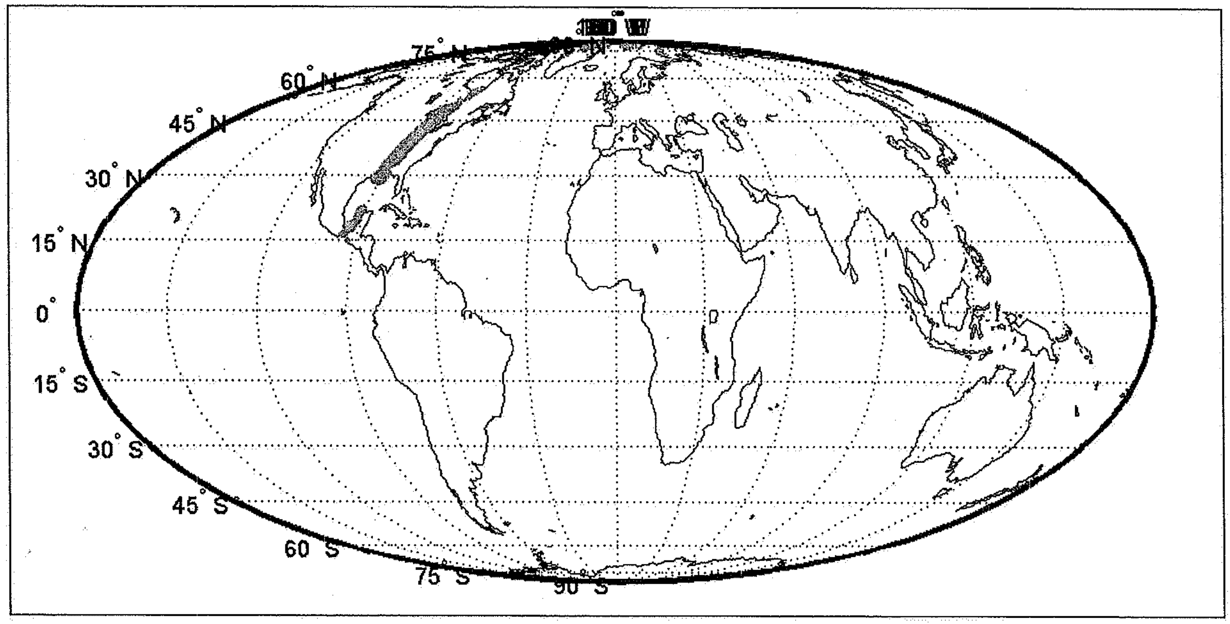
kwheeler@mail.arc.

MISR Rahman Model
Path 22

Ames Research Center

Path 22

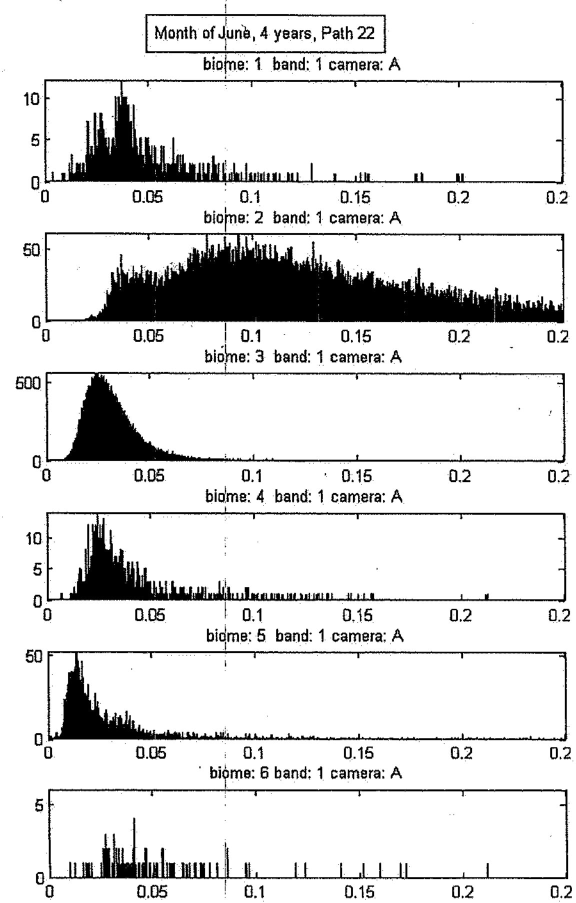
noFiles=60 totalBytes=673.7464

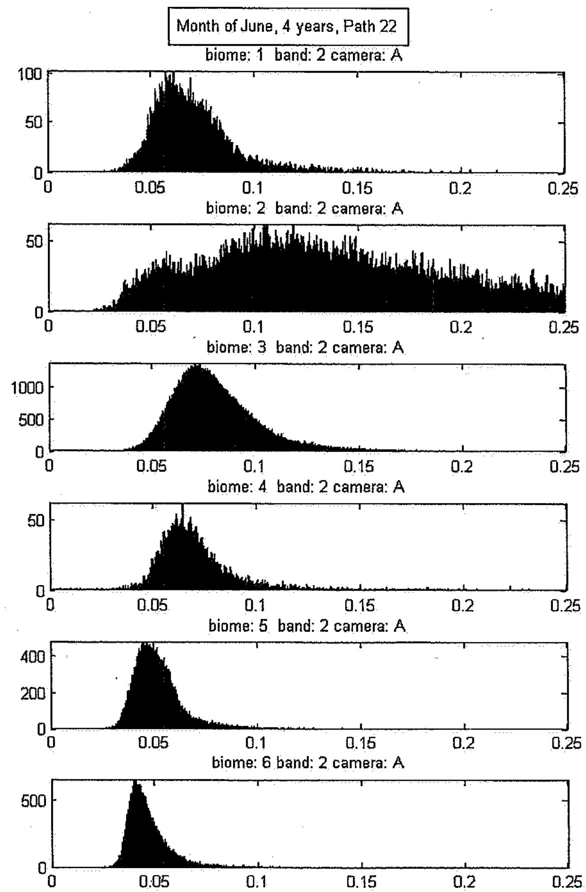


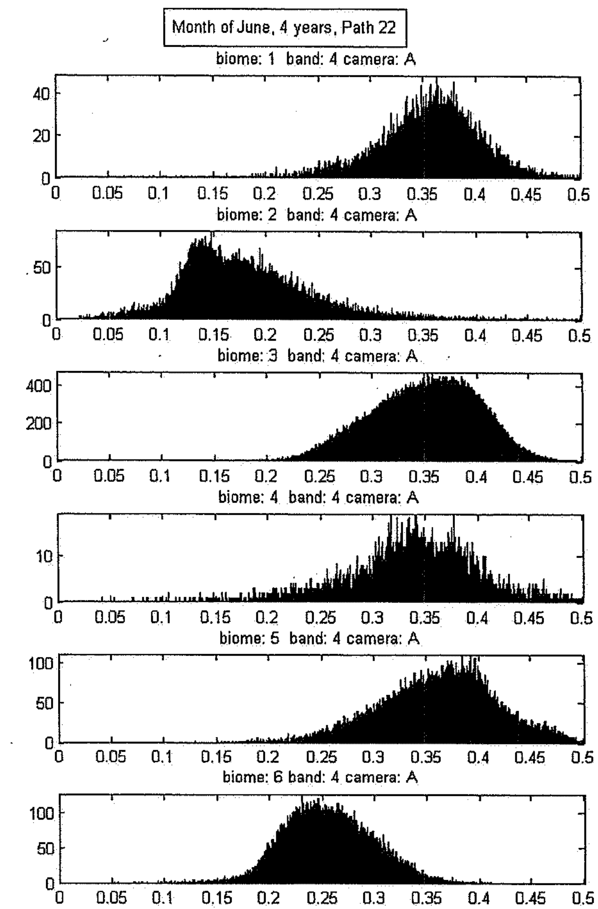
 kwheeler@mail.arc.nasa.gov

MISR Rahman Model Output

Camera A
Month of June, 2005 – 2001, Path 22







Ames
Research
Center

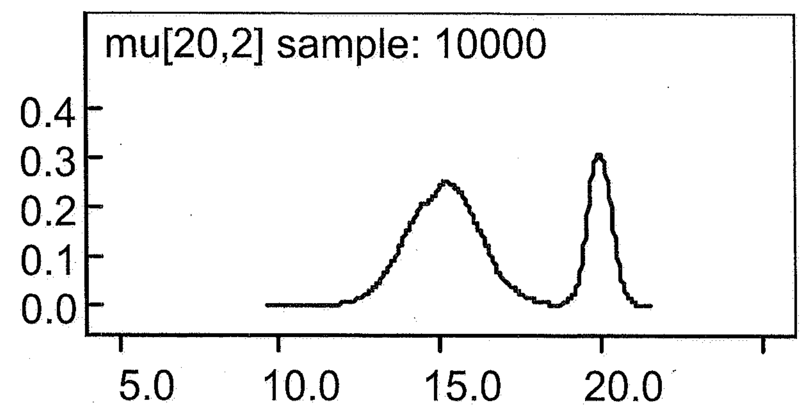
MISR Rahman Model
Dirichlet Processes

Goal: Analytic expression for each histogram no matter the shape.

Want to specify non-Gaussian distributions in terms of a simple mixture model:

$$p(\vec{x}) = \sum_{i=0}^{N} w_i \mathrm{N}(\vec{x}, \vec{\mu}_i, \Sigma_i)$$

Can easily determine the number of mixtures, the means, and the variances via using Dirichlet process mixtures (uses Gibbs sampling). We have shown this to work well with MODIS data.



Conclusion

- Introduced Bayesian adaptive sampling for solving biomass estimation.
- Introduced characterization of MISR Rahman model parameters conditioned upon MODIS landcover.
- Introduced rigorous non-parametric Bayesian approach to analytic mixture model determination.
- Introduced unique U.S. asset for science product validation and verification.

kwheeler@mail.arc.nasa.gov